

$$(3b) \quad Z_1 = \begin{bmatrix} 12 & 4 \\ 2 & 16 \\ 13 & 2 \\ -16 & 1 \end{bmatrix} \quad A_1 = \begin{bmatrix} 12 & 4 \\ 2 & 16 \\ 13 & 2 \\ 0 & 1 \end{bmatrix}$$

$$Z_{\text{out}} = \begin{bmatrix} 64 & 0 \\ -16 & -80 \end{bmatrix} \quad A_{\text{out}} = \begin{bmatrix} 64 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\frac{\partial \text{Loss}}{\partial W_{\text{out}}} = \begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix} \cdot \text{relu_dr}(Z_{\text{out}}) \cdot A_1^T$$

$\nwarrow \text{mask}$

$$= \begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 12 & 2 & 13 & 0 \\ 4 & 16 & 2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 12 & 2 & 13 & 0 \\ 4 & 16 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 36 & 6 & 39 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$\nwarrow \text{mask}$

$$\frac{\partial \text{Loss}}{\partial b_{\text{out}}} = \begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix} \cdot \text{relu_dr}(Z_{\text{out}}) \cdot [1 \ 1]^T$$

$$= \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$